|  |  |
| --- | --- |
|  | **KHOA CÔNG NGHỆ THÔNG TIN**  BẢN TÓM TẮT ĐỀ CƯƠNG ĐỒ ÁN TỐT NGHIỆP |

**TÊN ĐỀ TÀI: ÁP DỤNG THUẬT TOÁN PHÂN CỤM ĐỐI VỚI PHÂN LOẠI MỨC ĐỘ PHÁT TRIỂN CỦA CÁC QUỐC GIA**

*Sinh viên thực hiện*:

*Lớp*:

*Mã sinh viên:*

*Giáo viên hướng dẫn*:

**TÓM TẮT ĐỀ TÀI**

Hiện nay, mức độ phát triển của các quốc gia được đánh giá dựa trên nhiều yếu tố cơ bản như GDP, dân số v.v... Điều này giúp chúng ta có thể tiếp cận và đánh giá mức độ phát triển của các quốc gia. Tuy nhiên vấn đề được đề cập đến nhiều nhất là làm sao có thể phân loại được các tiêu chí để đánh giá khi mà số liệu về các yếu tố đó lại quá lớn và chúng ta không thể thu thập thông tin của từng người dân trong mỗi quốc gia đó.

Mục tiêu của các thuật toán phân cụm là từ tập dữ liệu khổng lồ, làm sao chúng ta biết có những nhóm dữ liệu đặc trưng nào trong đó? Từng dữ liệu trong đó thuộc vào nhóm nào? Đó là cái mà thuật toán phân cụm của chúng ta cần đi tìm câu trả lời. Đề tài này sẽ tập trung vào tìm hiểu Thuật toán phân cụm K-Means (K-Measn Clustering) và Thuật toán phân cụm phân cấp (Hierarchial Clustering) . Từ đó xây dựng bài toán “phân loại mức độ phát triển của các quốc gia” dựa trên Thuật toán phân cụm này để áp dụng với đối tượng là các quốc gia trên thế giới.

**CÁC MỤC TIÊU CHÍNH**

* Tìm hiểu về Thuật toán phân cụm K-Means & Hierarchial Clustering
* Tìm hiểu về bài toán phân loại mức độ phát triển của các quốc gia
* Xây dựng, lựa chọn Thuật toán phân cụm để khai phá dữ liệu và thống kê các chỉ số phát triển của các quốc gia.

**KẾT QUẢ DỰ KIẾN**

* Hiểu được về Thuật toán phân cụm K-Means & Hierarchial Clustering.
* Hiểu được bài toán phân loại mức độ phát triển của các quốc gia
* Dùng Python hoặc một số platform khác để phân loại bằng thuật toán phân cụm K-Means.